Санкт-Петербургский государственный университет

Математическое обеспечение и администрирование информационных систем
Технология программирования

Бушмелев Федор Витальевич

Профиль компетенций злоумышленника в имитации социоинженерных атак

Бакалаврская работа

Научный руководитель:
доц. каф. инф., к. пс. н., доц. Тулупьева Т. В.

Рецензент:
с.н.с. лаб. ТиМПИ СПИИРАН
к.ф.-м.н. Суворова А. В.

Санкт-Петербург
2018

SAINT-PETERSBURG STATE UNIVERSITY

Software and Administration of Information Systems
Technology of Programming

Bushmelev Fedor

Attacker's competencies profile social engineering attacks simulation

Bachelor's Thesis

Scientific supervisor:
Associate Prof., Computer Science Chair
Candidate of Psychology, Associate Prof. Tatyana Tulupyeva

Reviewer:
Senior Researcher, TICS Lab SPIIRAS
Candidate of Physico-Mathematical Science A. V. Suvorova

Saint-Petersburg
2018

Оглавление

Введение

Актуальность.

На сегодняшний день информационные технологии являются неотъемлемой частью повседневной жизни человека. Системы, что хранят, передают и обрабатывают информацию, используются повсеместно. Стремительный темп развития информационных технологий заставляет все больше внимания уделять вопросам информационной безопасности [4]. Причинами этого служит то, что за последние годы, в значительной мере увеличилось число атак на информационные системы [7]. Всё больше требуется средств и временных затрат для расследования и устранения их последствий [26].

Сегодня вопрос по обеспечению безопасности и конфиденциальности информации стоит наиболее остро. Большая часть исследований направлена на исследование программно-технических аспектов данной проблемы [20]. Надо отметить, что имеются очень серьезные наработки по этому вопросу, получены значимые результаты, но исследования продолжаются с не меньшей интенсивностью [6].

К сожалению, не смотря на все вышеупомянутые достижения, СМИ в изобилии рассказывают о самых разных и необычных инцидентах, связанных с нарушениями информационной безопасности [5, 19, 21, 23, 25]. Заметим, когда говорят о защищенности информационной системы, кибербезопасности или информационной защите, чаще всего имеют ввиду программно-техническую сторону вопроса, забывая, что пользователь также является одной из наиболее важных и уязвимых частей информационной системы и имеет непосредственное влияние на уровень ее защищенности. Таким образом, становится актуальной проблема по защите пользователей, и как результат – информации, от социоинженерных атак, т.е. атак, основной целью которых является персонал информационных систем. Недавняя статистика, представленая в [2], подтверждает актуальность проблематики социоинженерных атак. За прошедший год было зафиксировано более 53 000 инцидентов, связанных с информационной безопасностью, и 2216 нарушений. Для срав-

нения, за 2017 год [1] было зафиксировано 42 068 инцидентов, из них 1935 нарушений. Важно отметить, что 43% всех нарушений связаны с **социальной инженерией**, где большая часть (70%) — фишинг; 20% — непосредственное воздействие на жертву; и 10% — прочее. Наиболее подверженными отраслями являются: государственные и общественные организации, здравоохранение и образование. Другими словами, в первую попадают под удар учреждения, которые имеют дело с колоссальным количеством конфиденциальной информации. К сожалению, злоумышленники успешно получают эти данные [15, 23, 24, 25].Тем самым поднимается еще один важный вопрос — защищенность критичной информации в информационных системах.

Степень разработанности темы. Исследователями лаборатории теоретических и междисциплинарных проблем информатики Санкт-Петербургского института информатики и автоматизации
РАН (ТиМПИ СПИИРАН) была предложена модель социоинженерной атаки и разработан прототип программного комплекса, моделирующий социоинженерные атакующие воздействия на пользователей информационной системы [13, 14]. В основе лежит набор моделей: «критичные документы – информационная система – персонал – злоумышленник». Одной из наиболее проработанных является модель «персонал», чего нельзя сказать о моделях «критичные документы» и «злоумышленник». Но, даже не имея непосредственного контакта с предметом исследования, результаты уже вполне ощутимы, существуют значимые наработки, разработаны модели [11, 10] профиля компетенций злоумышленника и предложен подход к его построению для оценки защищенности информационной системы от социоинженерных атак [16].

Целью данной работы является автоматизация построения оценки степени защищенности категорий критичных документов в информационной системе. Иными словами, необходимо автоматизированно получить вероятностную оценку того, что злоумышленник с некоторым профилем компетенций получит доступ к критичным данным конкретного уровня конфиденциальности.

Для достижения цели были поставлены и решены следующие **задачи**:

- Изучить предметную область, источники по тематике исследования, описывающие подходы к автоматизированному анализу защищённости пользователей информационных систем от социоинженерных атак.

- Предложить подход к оценке защищённости критичных документов в информационной системе в рамках социоинженерных атак.

- Разработать вероятностную модель оценки степени защищённости критичных документов информационной системы от социоинженерных атак.

- Построить алгоритм оценки защищенности критичных документов, распределённых по степени критичности.

- Реализовать построенный алгоритмы в прототипе модуля комплекса программ.

Объектом исследования являются модель злоумышленника, в частности профиль компетенций и его ресурсная база, и критичные документы, находящихся в информационной системе.

Предметом исследования являются методы автоматизированной оценки защищенности категорий критичных документов от социоинженерных атак.

Научная новизна выпускной квалификационной работы заключается в том, что предложен новый подход к оценке защищённости критичных документов в информационной системе. Впервые предложена модель оценки степени защищённости критичных документов. Впервые реализован алгоритм оценки защищенности критичных документов, распределённых по степени критичности в прототипе модуля комплекса програм.

Теоретическая и практическая значимость. Представленная модель позволит оценивать уровень защищенности категорий критичной информации, имеющейся в информационной системе, на основе данных о пользователях, данных об информационной системе и пред-

ставлениях о злоумшленнике. Разработанный подход к построению профилей злоумыленников позволяет судить не только о защищенности критичных документов, но и открывает большие возможности для дальнейших разработок. Результаты представленной работы могут быть очень полезны работодателям, специалистам по безопасности, логистике и по работе с персоналом. Все предложенные подходы, модели, алгоритмы и программные решения поспособствуют дальнейшим исследованиям посвященным прогнозированию угроз и оценке защищенности от социоинженерных атак.

Методология бакалаврской работы заключается в постановке и формализации задач, связанных с автоматизированными оценками защищенности категорий критичных документов информационной системы; описанием моделей и сущностей, используемых для оценки; разработке алгоритмов и методов, применяемых для вычисления этих оценок, и релизации предложенных теоретических выкладок в качестве прототипов программных модулей.

Методы. Для проведения исследования в рамках бакалаврской выпускной квалификационной работы были использованы подходы и методы таких областей знаний как теория вероятностей, теори графов и нейронных сетей в качестве теоретических выкладок. Так же использовались методы сравнения и анализа для набора пороговых функций. Для реалиации же практической части работы использовались методы объектно-ориентированного программирования. Программная реализация осуществлялась в среде разработки Intellij IDEA 2017 на языке программирования Java. Для построения графических интерфейсов использовалась библиотека Java Swing.

Положения, выносимые на защиту:

- Подход к оценке защищённости критичных документов в информационной системе в рамках социоинженерных атак.
- Вероятностную модель оценки степени защищённости критичных документов информационной системы от социоинженерных атак.
- Реализация алгоритма оценки защищенности критичных доку-

ментов, распределённых по степени критичности в прототипе модуля комплекса программ.

высокая степень **достоверности** результатов данной выпускной квалификационной работы подтверждается глубоким и всесторонним анализом тематик, связанных с исследованиями по информационной безопасности и социоинженерным атакам, корректным применениеми математических методов и практик, подтверждается согласованностью полученных результатов, а также их успешной апробацией на международных и российских научных конференциях и публикациями в российский и международных изданиях.

Апробация результатов исследования. Результаты данной работы были представлены на следующих научных конференциях:

- Юбилейная X Санкт-Петербургская межрегиональная конференция «Информационная безопасность регионов России (ИБРР-2017)»
- VII-я Всероссийская научно-практическая конференция «Нечеткие системы, мягкие вычисления и интеллектуальные технологии»

Результаты данной выпускной квалификационной работы используются в рамках научно-исследовательского проекта, поддержанного грантами РФФИ № 18-37-00323 «Социоинженерные атаки в корпоративных информационных системах: подходы, методы и алгоритмы выявления наиболее вероятных траекторий» и № 18-37-00323 — «Методы анализа устойчивости структуры социальных связей пользователей информационной системы к социоинженерным атакующим воздействиям злоумышленника на основе применения генетических алгоритмов».

Публикации. По теме данной бакалаврской работы было сделано 4 публикации: 2 из которых индексируются РИНЦ [17, 16], 1 работа принята к публикации, индексируемая изданиями Scopus / WoS.

Благодарности. Данная выпускная квалификационная работа бакалавра содержит материалы исследований, выполняемых в рамках государственного задания СПИИРАН № 0073-2018-0001, а также поддержанных грантами РФФИ: № 18-37-00323 — «Социоинженерные атаки

в корпоративных информационных системах: подходы, методы и алгоритмы выявления наиболее вероятных траекторий» и № 18-37-00323 — «Методы анализа устойчивости структуры социальных связей пользователей информационной системы к социоинженерным атакующим воздействиям злоумышленника на основе применения генетических алгоритмов»

Структура работы. Текст данной работы состоит из введения, 4 глав, заключения, списка используемой литературы приложения со списком терминов. Общий объем — 37 страниц.

В 1 главе строится представление о предметной области, обосновывается выбор и постановка цели и задач.

Во 2 главе описывается научная база и различные средства, послужившие основой для проведения данной работы. Приведен обзор используемых технических средств.

В 3 главе описываются разработанные подходы, модели и алгортмы к оценке защищенности категорий критичных документов информационной системы от социоинженерных атак. Представлены основные теоретические результаты выпускной квалификационной работы.

В 4 главе предоставляется описание разработанных прототипов программных средств.

1. Описание предметной области

Данная глава посвящена проблеме подвержености информационных систем социоинженерным атакам. Обосновывается необходимоть оценки защищенности категорий критических документов. В главе представлены актуальность, обоснование целей и задач, поставленных в данной работе.

1.1. Актуальность

На сегодняшний день вопрос по обеспечению безопасности в информационных системах стоит наиболее остро. Не смотря на всё изобилие разного рода решений и технологий, направленных на защиту информации, количество инцедентов не сократилось. Это обосновано тем, что упомянутые подходы направлены на защиту программно-технических аспетов информационных систем. Таким образом на первый план выходит пробелема связанная с обеспечением безопасности от атак направленных на пользователей или социоинженерных атак. Чем и подтверждается актуальность данной тематики. Уже существуют программные решения [14] позволяющие моделировать многоходовые социоинженерные атаки, выявлять уязвимости и некоторые оссобенности пользователей информационной системы и многое другое. Однако, не смотря на то, что данный программный косплекс решает широкий класс задач, назвать его полным никак нельзя. Есть достаточное количество направлений для развития и уточнения этого комплекса. Ввиду того, что злоумышленник при социоинженерной атаке может преследовать различные цели, например, критичные документы, имеется необходимость в анализе их защищенности при социоинженерной атаке, чего пока не реализовано. Также было бы полезно иметь представление о возможностях злоумышленника, ресурсах. Реализация данных решений позволит повысить защищенность критических документов и самих информационных систем, сведя к минимуму возможный ущерб организации от действий социального инженера, и как следствие снижение количества успешных атак подобного рода. Данная работа является частью общего

исследования, цель которого заключается в автоматизации анализа защищённости информационных систем от социоинженерных атак. Цель, достигаемая в этой выпускной квалификационной работе, состоит в автоматизации процессов анализа защищенности категорий критических документов.

1.2. Цели и задачи

Ввиду того, что есть потребность дополнении упомянутого ранее программного решения для большей автоматизации оценок защищенности, было принято решение предложить несколько подходов и моделей к анализу защищенности от социоинженерной атаки и разработать два прототипа модулей, для автотизированной оценке защищенности конфиденциальной информации.

Целью данной работы является автоматизация построения оценки степени защищенности категорий критичных документов в информационной системе. Иными словами, необходимо автоматизированно получить вероятностную оценку того, что злоумышленник с некоторым профилем компетенций получит доступ к критичным данным конкретного уровня конфиденциальности.

Для достижения цели необходимо решить следующие **задачи**:

- Изучить предметную область, источники по тематике исследования, описывающие подходы к автоматизированному анализу защищённости пользователей информационных систем от социоинженерных атак.
- Предложить подход к оценке защищённости критичных документов в информационной системе в рамках социоинженерных атак.
- Разработать вероятностную модель оценки степени защищённости критичных документов информационной системы от социоинженерных атак.
- Построить алгоритм оценки защищенности критичных документов, распределённых по степени критичности.

- Реализовать построенный алгоритмы в прототипе модуля комплекса программ.

2. Используемые подходы и решения

В данной главе приводтся обзор работы, ставшими основой для написания данной выпускной квалификационной работы, описываются используемые в работе программные средства.

2.1. Исследования, взятые за основу работы

Заделом для данной работы являются исследования проводимые в лаборатории ТиМПИ СПИИРАН. Подробные результаты их исследований представлены в труде [14]. В нем собраны разнообразные модели, подходы, алгоритмы для оценки успехов социоинженерной атаки, многоходовой социоинженерной атаки, имитации атаки основанной на деревьях атак, предложена модель «злоумышленник» и многое другое. Однако, в данном исследовании не было упоминаний ни о подходах к оценке защищенности данных в информационной системе, ни об ее автоматизации. Также в данной работе хоть и была представлена модель «злоумышленник», но давала лишь о злоумышленнике как некой сущности. Сейчас же существуют модели для оценки защищенности конфиденциальньой информации, модель «злоумышленник» стала более проработанной [11]. К сожалению, все также нет удобного инструмента позволяющего говорить о степени защищенности документов в информационной системе. Тоже самое относится и к получению портрета злоумышленника, позволяющее также оценить защищеность данных от социоинженерных атак.

2.2. Обзор существующих моделей вероятностной оценки социоинженерной атаки с целью получения доступа к критичному документу

В [9] были предложены модели вероятностной оценки успешно подействовать j-м социоиненерным воздействием на i-ую уязвимость пользователя. Также в [9] был предложен подход к оценке успеха социо-

инженерного воздействия злоумышленника на пользователя с целью получения доступа к критичному документу некоторой категории критичности. Данный подход использует адаптированную модель Белла–Тревино [3]. Модель Белла–Тревино — часто используемый в биоинспирированных вычислениях комплекс моделей, которые связывают оценку риска с числом эпизодов рискованного поведения.

2.3. Описание используемых программных средств

Ввиду того, что вышеупомянутый прототип программного комплекса может использовать различные существующие программные решения, в том числе и сторонние, а также принимая в расчет возможность внедрения в данный комплекс разрабатываемых в данной работе модулей, было принято решение, что их реализация будет проходить на языке программирования Java, как программный комплекс. Также этот выбор обоснован тем, что модули в дальнейшем могут расширяться и для решения потребуется использование сторонних средств. Список средств используемых при разрботке:

- Язык разработки Java 8 — мощной кроссплатформенный объектоориентированный язык программирования .
- Java Swing — библиотека для создания графического интерфейса для программ на языке Java.
- Среда разработки Intellij IDEA 2017 – интегрированная среда разработки на Java;

3. Оценка защищенности критичных документов

В этой главе представлены основные теоретические результаты выпускной квалификационной работы бакалавра. Она посвящена разработке подходов, моделей, алгоритмов и программных модулей, для автоматизированной оценки защищенности критических документов и построения предстравления о профиле компетенций возмоного злоумышленника. Также предложена модель учета ресурсов в социоинженерной атаке.

3.1. Подходы к оценке критичности документа

Как было отмечено ранее, важным аспектом в анализе защищенности информационной системы является возможность вывода оценок защищенности критичных документов, хранимой в данной системе. Очевидно, что определение критичности документов — отнесение их к той или иной категории по уровню критичности — может быть осуществлено различными способами. Одним из часто используемых методов выделения категорий критичности информации, хранящейся в системе, является финансовая оценка возможного ущерба, который может быть нанесен при получении несанкционированного доступа злоумышленника к некоторому критичному документу.

Представим, что имеется некая информационная система, будем рассматривать ту её часть, которая состоит из пользователей и наборов доступных критичных документов для каждого пользователя системы. Таким образом на Рис.1 приведен пример описанной системы, где у 3 пользователей имеется доступ к критичным документам (обозначается как «КД», далее следует имя пользователя с доступом к этому документу и номер документа). Важно отметить, что один документ из перечня "Пользователя А" может соответсвовать некоторому документу "Пользователя В".

Пользователи в соответствии с моделью «персонал» имеют один

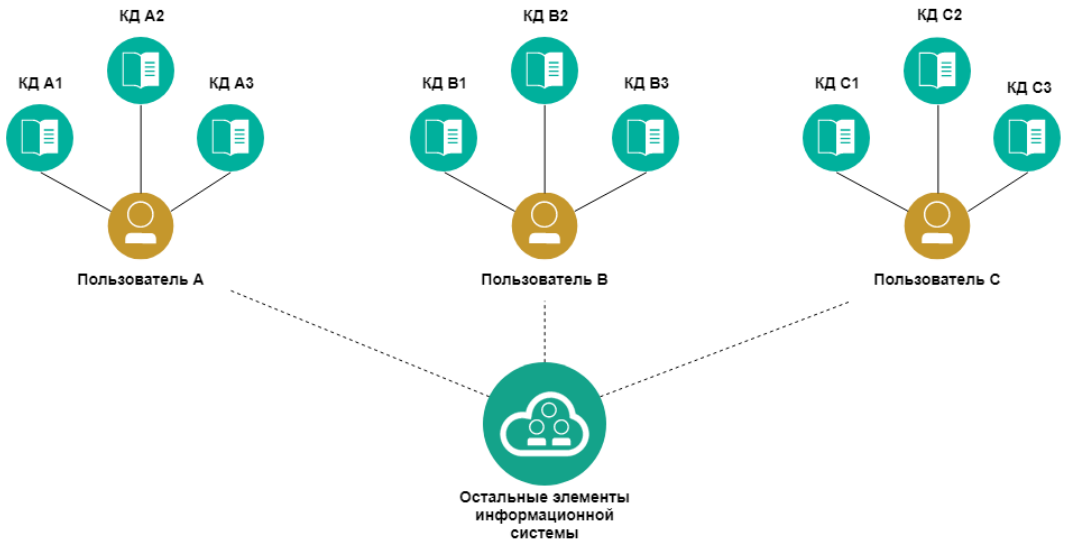

Рис. 1: Пример информационной системы с выделением пользователей, имеющих доступ к набору критичных документов

наиболее важный сейчас аспект — профиль уязвимостей пользователя, который задается парой «уязвимость» – «выраженность уязвимости». Аналогично, в модели «злоумышленник» можно выделить «атакующее воздействие» – «степень владения атакующим воздействием».

Теперь приступим к разработке модели и алгоритма по оценке степени защищенности категорий критичных документов.

3.2. Учет ресурсов в оценках защищенности в задаче социоинженерных атак

Ранее рассматривалась ситуация, когда доступ к документу атакующий старался получить за счет своих компетенций. Сейчас это становится уже недостаточным, необходимо получить представление о роли ресурсов в оценках социоинженерных атак. Обратимся еще раз к модели «злоумышленник», а именно к ресурсной базе злоумышленника. Будем представлять эту базу в виде вектора, обозначив w_s за текущее значение s-го ресурса. Также надо упомянуть, что для эффективного влияния на i-ю уязвимость пользователя, злоумышленнику необходимо затратить некоторое значение каждого ресурса. Обозначим r_{is} — количество s-го ресурса, чтобы повлиять на i-ую уязвимоть.

Исследуем поведение функции $R(w_s, r_{is})$, обозначающей вероятность влияния на полученную ранее вероятность, в зависимости от имеющихся у злоумышленника и требуемых для воздействия ресурсов. Очевидно, что при $\forall r_{is}$ и $w_s \to max$ значение $R(w_s, r_{is}) = 1$, а при $\forall r_{is}$ и $w_s = 0$ значение $R(w_s, r_{is}) = 0$. Иными словами, функция $R(w_s, r_{is})$ может быть представлена пороговой функцией [18]. Далее рассмтрим несколько видов пороговых функций. Важно отметить, что рассмотренные варианты не являются строго фиксировнными, возможны и другие способы задать данную функцию. Обозначать за $R_s^i(w_s, r_{is})$ будем пороговую функцию показывающую влияние s-го ресурса на i-ую узявисомть пользователя; $R^i(w_s, r_{is})$ – аналогично, только по совокупности ресурсов атакующего.

Пороговая передаточная функция или **функция Хевисайда** является одним из классических примеров простой пороговой функции. Пока злоумышленник имеет достаточное количество ресурсов, чтобы покрыть потребности пользователя, доступ к документам будем считать получен:

$$R_s^i(w_s, r_{is}) = \begin{cases} 1, & 0 \leqslant r_{is} \leqslant w_s; \\ 0, & 0 \leqslant w_s < r_{is}. \end{cases} \tag{1}$$

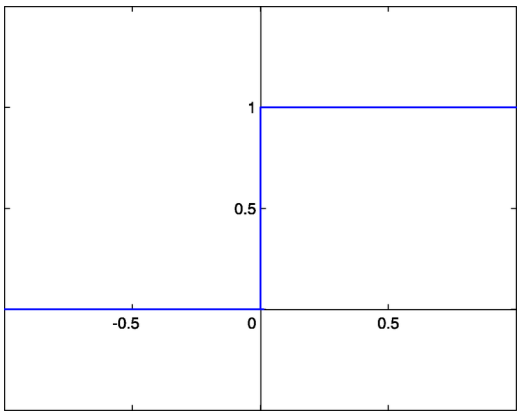

Рис. 2: Пример функции Хаусдорфа. Источник изображения [18]

При анализе функции $R(w_s, r_{is})$ мы рассматривали только крайние случаи. Для ситуации когда $|w_s - r_{is}| < \epsilon$, где ϵ – малая окрестность r_{is}, всё становится не так очевидно, за исключением самых простейших

случаев, например, рассмотренная ранее функция Хаусдорфа. Дальше речь пойдет о более сложных функциях, имеющих параметр α. Он отвечает за то, как будет вести себя функция в окрестности r_{is}. При $\alpha \to 0$ график пороговой функции будет становиться более пологим, при $\alpha \to \infty$ график вырождается в функцию Хаусдорфа.

Линейная передаточная функция очень похожа по поведению на функцию Хаусдорфа, за исключением того, что в окрестности r_{is} вероятность передечи документа линейно начинает возрастать в зоне действия параметра α.

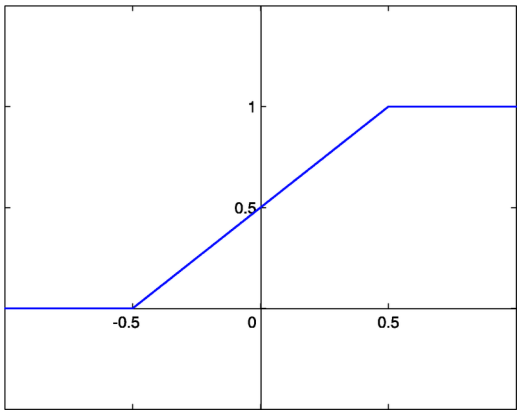

Рис. 3: Пример линейной передаточной функции. Источник изображения [18]

Экспоненциальная функция передеачи задается следущим образом:

$$R_s^i(w_s, r_{is}) = \begin{cases} 1 - e^{-\alpha \cdot (w_s - r_{is})}, & 0 \leqslant r_{is} < w_s; \\ 0, & 0 \leqslant w_s \leqslant r_{is}. \end{cases} \tag{2}$$

Логистическая функция Она единственная из всех вышеперечисленных имеет наибольшее практическое применение в смежных тематиках. Задается следубщим образом:

$$f(x) = \frac{1}{1 + e^{-\alpha \cdot x)}}; \tag{3}$$

Применяя в терминах рассматриваемой задачи, получим влияние s-го ресурса на успех:

$$R_s^i(w_s, r_{is}) = \frac{1}{1 + e^{-\alpha \cdot (w_s - r_{is})}} \tag{4}$$

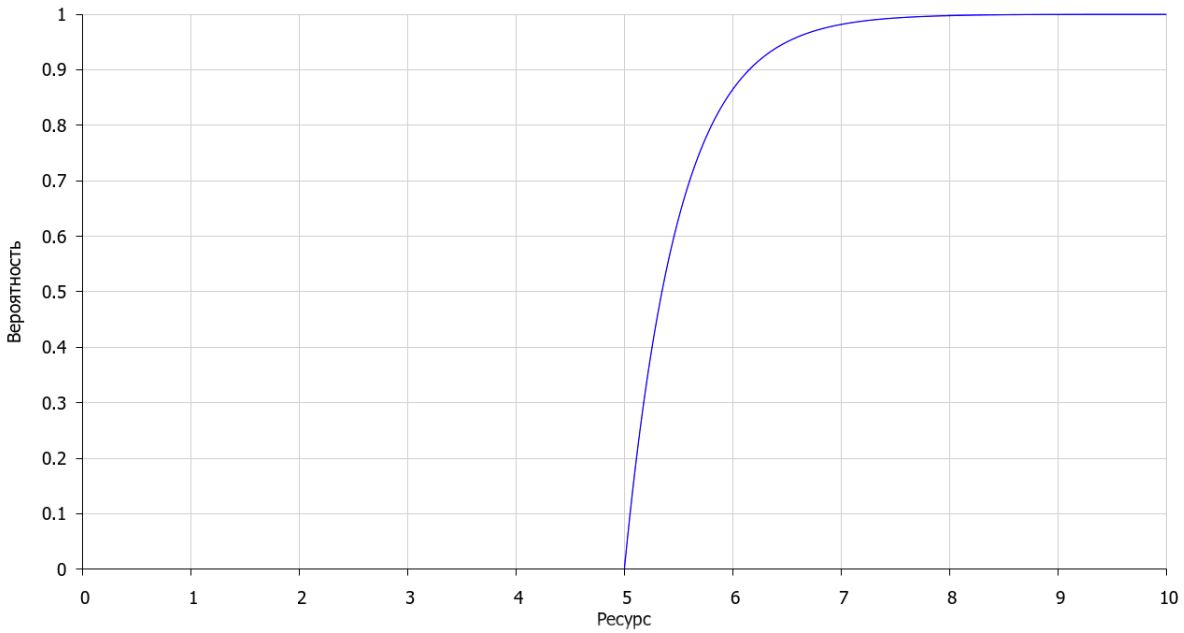

Рис. 4: Пример экспоненциальной передаточной функции

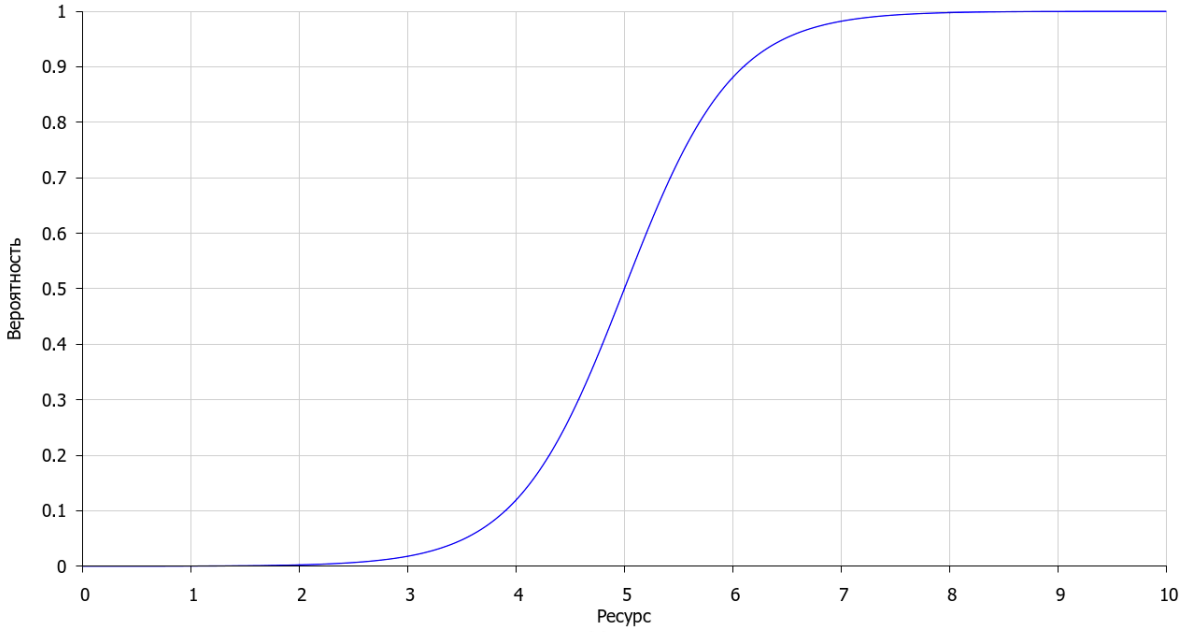

Рис. 5: Пример логистической функции

Влияние совокупности ресурсов будет выглядеть следущим образом:

$$R^i(w_s, r_{is}) = \prod_{n=1}^{m}\left(\frac{1}{1+e^{-\alpha\cdot(w_s - r_{is})}}\right); \tag{5}$$

где m — размерность вектора ресурсов.

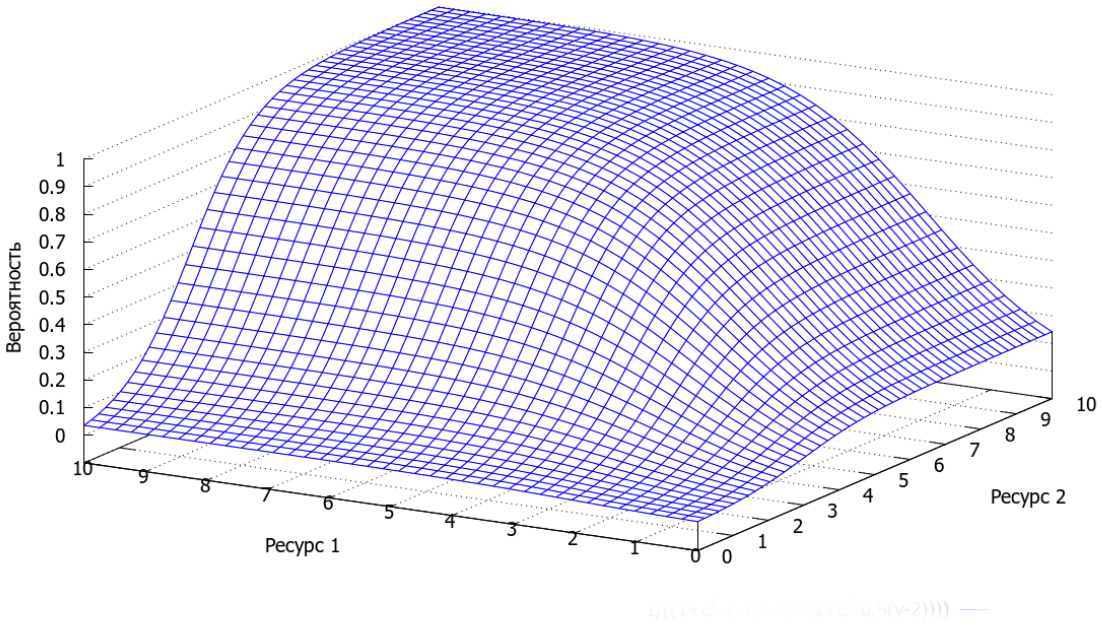

Рис. 6: Пример: график логистической функции для набора из 2 ресурсов

По итогу примененяя логичстическую функцию в вероятности получим:

$$P = 1 - \prod_{n=1}^{n}(1 - \overline{p_{ij}}), \tag{6}$$

где $\overline{p_{ij}} = p_{ij} \cdot R^i(w_s, r_{is}) = p_{ij} \cdot \prod_{n=1}^{m}(\frac{1}{1+e^{-\alpha\cdot(w_s - r_{is})}})$;

Остался нерешенным вопрос связанный с природой α, хотелось понять как ее задавть и отчего она зависит. Во всех рассмотренных нетривиальных случаях пороговой функции α понималась как константа задающая *кручение*. К сожалению, сейчас однозначно сказать о том, что это за параметр и как он задается нельзя. Это требует отдельного серьезного и обстотельного исследования. Но, для внесения некоторой определенности сделаем следующее предложение: вполне вероятно, что данный параметр зависит от личностных особенностей, также небезосновательным будет предположение, что злоумыленик может с помощью социоинженерного воздействия повлиять на α. Предлага-

ется говорить, что α показывает отношение, разницу в возможностях злоумышленника и пользователя, т.е. во сколько раз злоумышленник лучше владеет атакующим воздействием, чем пользователь сопротивляется уязвимости и каково влияние одного на другое. Формализуем предложени:

$$\alpha_{ij} = \frac{S(A_j)}{1 - D(V_i)} \cdot \ q_{ji}. \tag{7}$$

3.3. Профиль компетенций злоумышленника как средство к оценке защищенности критичных документов

Для оптимальной оценки защищенности критичной информации предлагается также принимать в рассмотрение не только степень выраженности уязвимостей пользователей информационных систем, но и вероятностные оценки сил и средств, которыми мог бы обладать злоумышленник. Будем считать, что социоинженерная атака была успешно проведена, и доступ к некоторому документу злоумышленником получен. Необходимо установить какими минимально возможными ресурсами мог бы обладать такой злоумышленник. В рассматриваемой задаче будет использовано представление информационной системы, приведенное в [12]. На Рис.7 приведен пример графа социальных связей пользователей информационной системы, где узлы — это пользователи, обладающие всеми характеристиками модели «персонал», в том числе моделью профиля уязвимостей пользователя, а дуги графа — это вероятности перехода социоинженерной атаки от пользователя к пользователю.

Описание подхода к оценке

Необходимо перестроить граф для визуализации предлагаемого подхода. Для этого граф преобразуется к корневому, корнем которого будет являться критичный документ, а сыновьями пользователи, имеющие к нему доступ, а внуками, пользователи, связанные с помощью дуг и т.д. Например, пусть пользователи 1,4 и 8 имеют доступ к какому-то

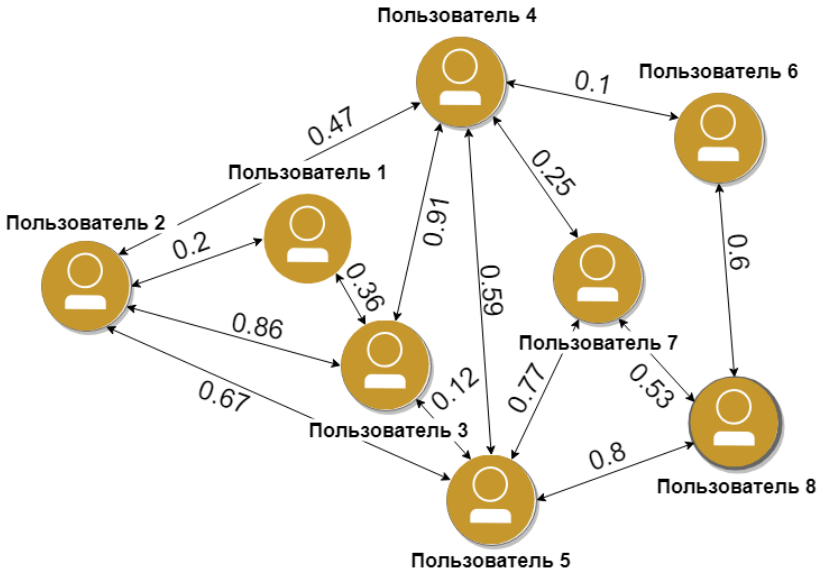

Рис. 7: Пример: граф межличностных связей персонала информационной системы

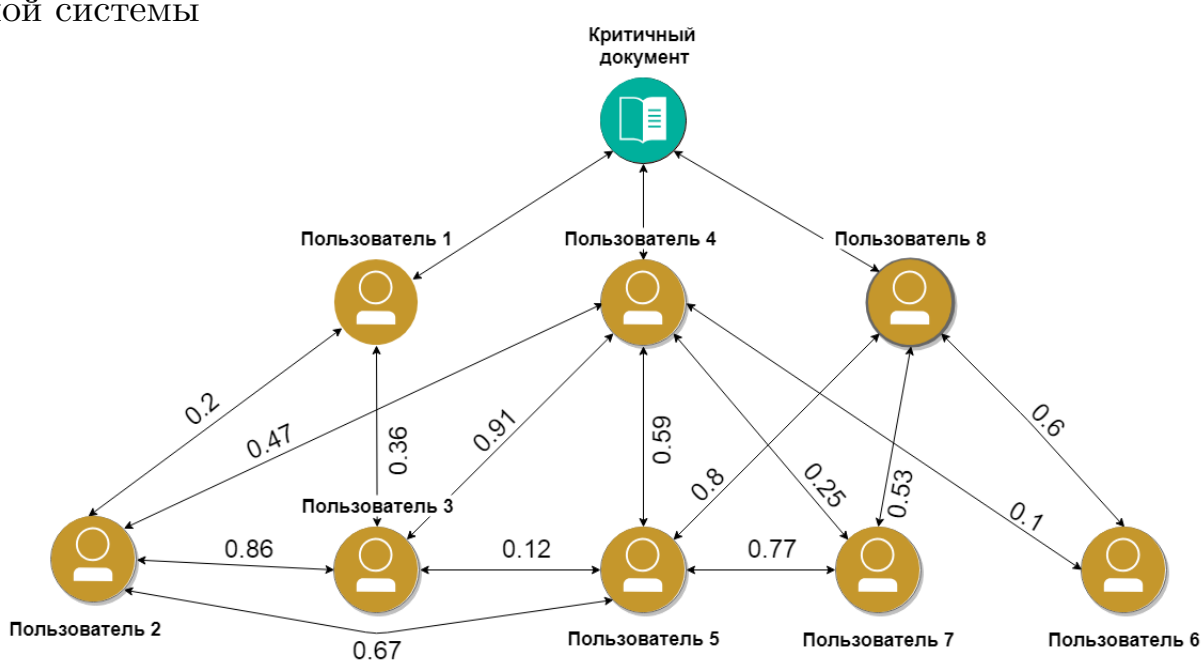

Рис. 8: Пример: преобразованный граф межличностных связей персонала с выделением некоторого критичного документа

критичному документу. Тогда визуальное отображение примера графа может быть представленно в виде Рис.8.

Опишем алгоритм предлагаемого метода. Установив метку текущего положения в корень, будем двигаться вниз, вычисляя возможные точки входа для атаки. Для начала необходимо рассмотреть всех соседей корневого элемента на некотором расстоянии s от него, где $s \in [0, V-1], V$

— мощность множества вершин графа. Достигнув некоторой вершины, высчитывается вероятностные оценки параметров профиля компетенций злоумышленника.

Формализация алгоритма

Обнаружив вершины, связанные с корнем, непосредственно формируется запись $\left\langle id_r^s, \{A_j, S(A_j)\}_{j=1}^m, p_r^s \right\rangle$, где id_r^s — уникальное имя злоумышленника, а r — последовательность вершин длины $s+1$, $\{A_j, S(A_j)\}_{j=1}^m$ — профиль компетенций злоумышленника, $S(A_j)$ — степень владения j-м атакующим действием, а p_r^s — вероятность достижения данной конфигурации степеней владения атакующими действиями последовательности r . Для описания метода, предполагаencz, что пользователь уже успешно атакован злоумышленником, и им уже получен доступ к критичной информации, которой обладает пользователь. То есть, не умаляя общности будем считать, что вероятность этого события 1. Тогда, имея профиль уязвимостей пользователя – из информационной системы, информацию о влиянии атакующих действий на уязвимости [14] и прочие сведения, необходимо будет решить экстремальную задачу и найти $S(A_j)$, иными словами: При $P_i = 1$, необходимо найти такие значения $\{S(A_j)\}_{j=1}^m \to min$, где $P_i = 1$ — вероятность того, что i-й пользователь был успешно атакован.

Замечание. Описывается ситуация, когда оценка идет якобы постфактум, когда критичная информация была получена и справедливо высказывание, что $P_i = 1$. Но, к сожалению, таким решением значительно суживается представление о профиле компетенций потенциальных атакующих. Для получения более точных результатов необходимо выставить для P_i подходящее значение, т.к. при учете менее вероятных исходов идет увеличение класса возможных злоумышленников, но также идет увеличение шума. Рис.9 служит более наглядным примером, где А — профили злоумышленников получаемые при $P_i = 1$; Б — например, при $P_i = 0.75$; и В — при $P_i = 0.5$. Это также привнесет изменения в алгоритм, добавив итерации для получения записей для каждого такого $P_i^k = \{P_i | P_i = 1 - t \cdot h, t \in [0, k]\}$, где h — шаг. После получения наборов записей для всех P_i^k, комбинирование и сепарация

полученных результатов.

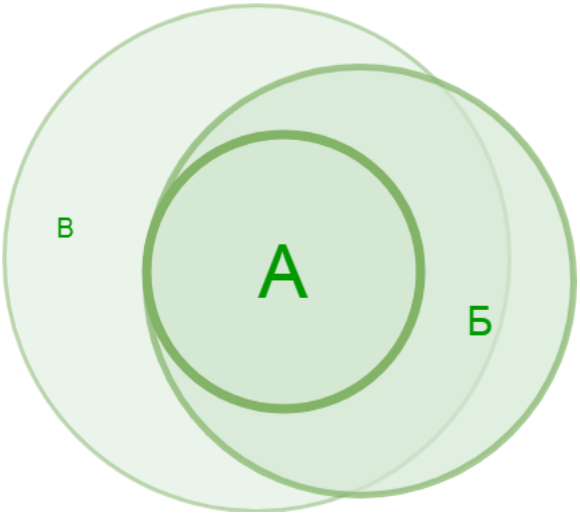

Рис. 9: Пример: диаграмма профилей компетенций злоумышленников при разных значениях постериорной вероятности

Важно отметить, что подход может иметь несколько вариаций для расчета p_r^s и $S(A_j)$ в зависимости от условий организации атаки. Не умаляя общности рассмотрим следующее: злоумышленник на всем пути, успешно атакует каждого пользователя по цепочки. Напомним, мы идем от корня к листьям, от документа к первому вхождению. При переходе по дуге приходится вести расчет $S(A_j)$ для каждого пользователя, и модифицировать значение с учетом ранее полученной выраженности компетенции.

Частный случай:

Примером вариативности может послужить социоинженерная атака, передающаяся посредствам инсайдерской атаки. Мы по-прежнему считаем, что документ успешно захвачен злоумышленником, имеется некая цепочка из $s+1$ узлов, тогда $p_r^s = \prod_{t \in r} p_t$, где p_t — вероятность передать атаку от пользователя к пользователю, t — пара смежных вершин из цепочки r. Подобная вероятность получается в результате инсайдерской атаки, где злоумышленнику необходимо атаковать только одного пользователя, в цепочке его номер будет последним, и дальнейшее распространение атаки происходит без значимых затрат для злоумышленника. Вычисление $S(A_j)$ потребуется выполнить только один раз.

Агрегация результатов

Предлгается вести агрегацию записей с самого первого шага алгоритма. Можно выделить несколько свойств:

- Если при формировании записей с длиной цепочки записи длины $s+1$ вероятность p_r^s меньшее порогового значения, например, 0.05, тогда эта запись не учитывается.

- Ввиду того, что для длинных цепочек, например, длины 5 и более, вероятность p_r^s в большинстве своем становится ничтожно мала, то такие записи учитываться также не будут.

- Получение новой цепочки длины $s+2$ происходит путем рассмотрения всевозможных цепочек длины $s+1$, и перехода из текущей вершины в любую из смежных к ней, которая была не посещена ранее. Для всех новых цепочек создаются записи.

- После получения всех возможных цепочек, производится склеивание записей с одинаковыми конечными звеньями цепочек. И после производится агрегация по всем записям, ведущих к рассматриваемому критичному документу.

После завершения работы алгоритма, будет получена одна запись с вероятностными оценками профиля компетенций злоумышленника.

Также приходится иметь ввиду тот факт, что современные информационные системы могут насчитывать сотни, и даже тысячи сотрудников, а каждый шаг алгоритма требует существенных временных затрат, даже на оборудовании высокой вычислительной мощности, не говоря обо всем подходе в целом, с расчетом для многих критичных документов. Для частичного решения данной проблемы видится возможным несколько вариантов оптимизации:

- Предварительно использовать алгоритм обхода графа в глубину с некоторыми модификациями и описанными ранее ограничениями по длине и вероятности перехода. При обходе вычисляется профиль компетенций злоумышленника, и каждой дуге присваивается $\{A_j, \overline{S(A_j)}\}_{j=1}^m$, где $\overline{S(A_j)}$ — разница между наборами инцидентных этой дуге вершин. И уже при непосредственной работе

алгоритма можно получить результат посредствам простого суммирования.

- Использовать представления моделей «информационная система» и «персонал», где используется деление пользователей по контролируемым зонам. Поэтому, можно осуществлять распределенную обработку по таким зонам.

Таким образом получаем сокращение расходов на вычисление и вместе с тем, система остается масштабируемой. Хотелось бы отметить, что данный подход к оценке профиля компетенций злоумышленника имеет более широкое применение, не только в вопросах связанных с защищенностью критичной информации.

4. Реализация прототипов программных модулей для оценки защищенности категорий критических документов

Данная глава посвящена описанию разработки программных модулей для построения представления о профиле компетенций злоумышленника и получения оценки защищенности категорий критических документов.

4.1. Программный модуль для оценки защищенности категорий критичных документов

Данный программный модуль предназначен для автоматизированной оценки степеней защищенности категорий критичных документов. Графический интерфейс представлен на 10.

Данные, необходимые для работы модуля, получаемые извне:

- Данные модели «персонал».
- Данные модели «критичные документы».
- Набор норм для калькуляции.
- Модель организации хранения документов в информационной системе.

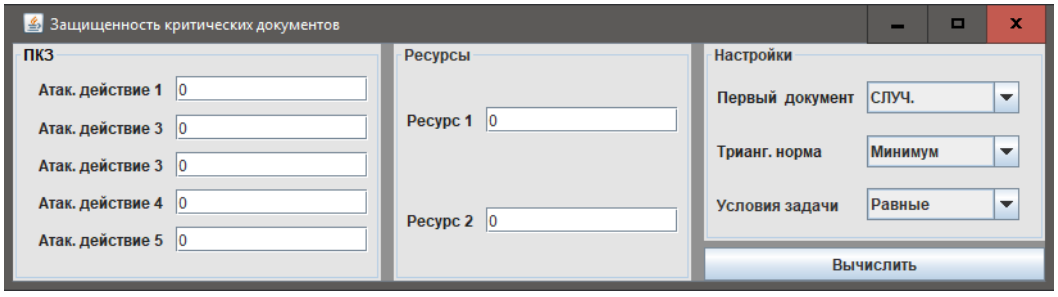

Рис. 10: Графический интерфейс для модуля оценки защищенности категорий критических документов

Система компонент модуля состоит из трех основных классов:

- **Input** — класс для отправки запросов, получение, предобработка данных извне и полученнойпутем ввода с графического интнрфейса и последущая передача в *ComputeProbability* необходимой информации. Сейчас используется чтение из файла, т.к. пока нет возможности внедрения модуля в программный комплекс.

- **ComputeProbability** — осуществялет расщет защищенности категорий критических документов, подрая работа алгоритма описана в Главе 3.

- **Output** — выводит реузьтат калькуляции, полученный в *ComputeProbability* на экран / в файл / ответ на запрос.

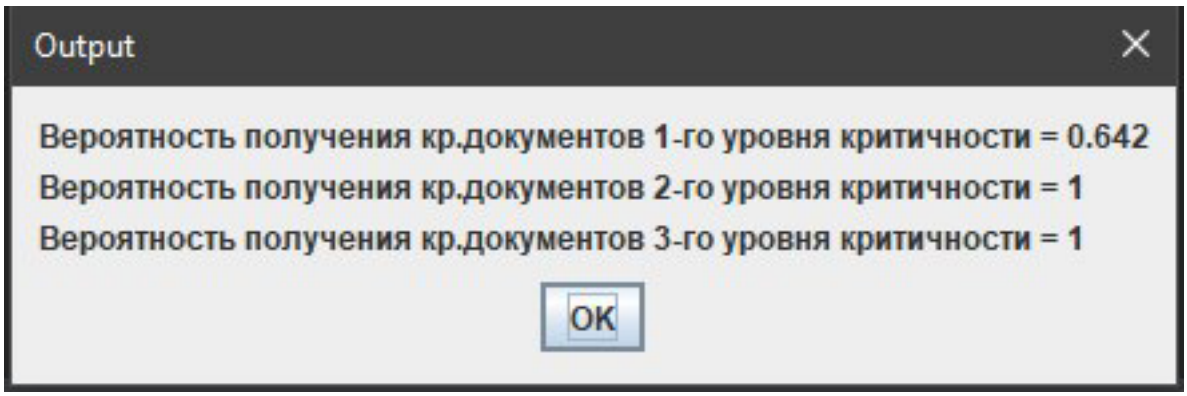

Рис. 11: Результат работы модуля

4.2. Программный модуль для построения профилей компетенций злоумышленника

Данный программный модуль предназначен для автоматизированной оценки степеней защищенности категорий критичных документов. Графический интерфейс представлен на 12.

Информация необходимая для работы модуля, аналогична той, используется в предыдущем пункте. Система компонент данного модуля схожа с предыдущим пунктом. Отметим важные особенности:

- **Competitions** — класс представлящий запись компетенций злоумышленника, описанной в Главе 3.

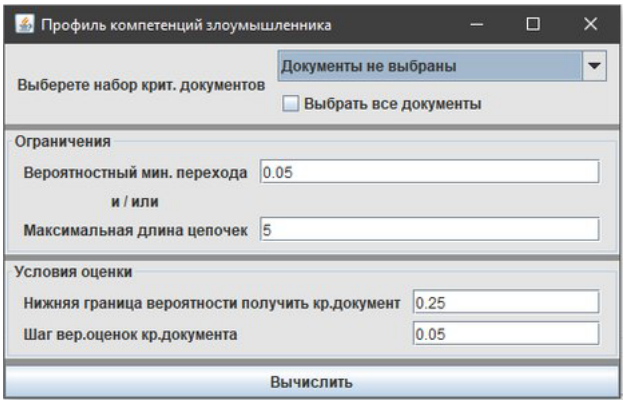

Рис. 12: Графический интерфейс для вычисления профиля возможного злоумышленника

- **constructGraph** — метод в *ComputeProbability*, отвечающий за формирование социального графа с ограничениями вокруг определенного документа.
- **computeComp** — метод в *ComputeProbability*, отвечающий за расчет и добавление/опускание записи с компетенциями в набор.
- **agrigateComp** — метод в *ComputeProbability*, обрабатывающий сначала записи с одинаковым конечным пользователем путем склеивания, далее агрегация с целью выявления профиля компетенций.

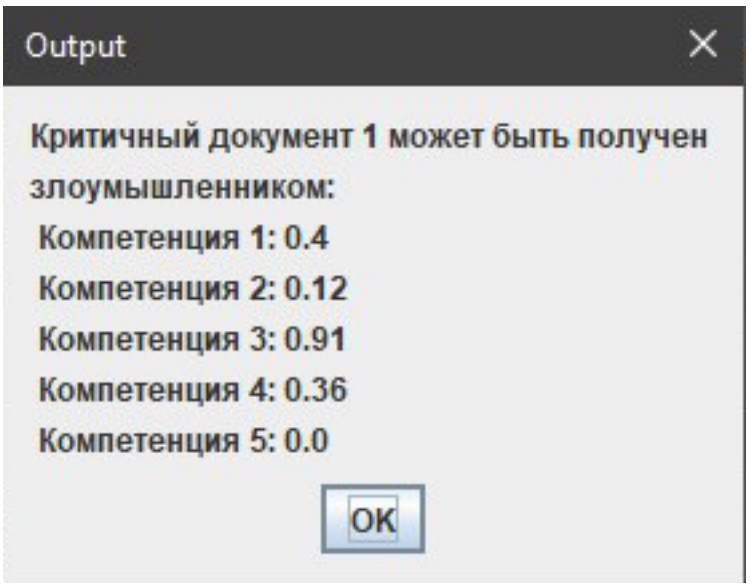

Рис. 13: Результат работы модуля

Заключение

Данная выпускная квалификационная работа бакалавра была посвящена автоматизации построения оценки степени защищенности категорий критичных документов в информационной системе. Что имеет большое значение для оценок защищенности информации, пользователей и всей информационной системы. Полученные результат формирует задел для дальнейшего развития решения вопросов безопасности связанных с социоинженерными атаками. Для достижения этой цели было выполнены следущие частные задачи:

- Проведен анализ предметной области, изучены источники по тематике исследования, описывающие подходы к автоматизированному анализу защищённости пользователей информационных систем от социоинженерных атак.
- Предложен подход к оценке защищённости критичных документов в информационной системе в рамках социоинженерных атак.
- Разработан вероятностная модель оценки степени защищённости критичных документов информационной системы от социоинженерных атак.
- Построен алгоритм оценки защищенности критичных документов, распределённых по степени критичности.
- Реализован построенный алгоритм в прототипе модуля комплекса программ.

По итогу, можно заключить, что все поставленные задачи выполнены и цель работы по автоматизации оценки защищенности категории критических документов была успешно достигнута. Результаты работы могут быть применены менеджерами компаний и специалистами по безопасности для наиболее эффективной и продуктивной работы коллектива функционирования организации. Результаты данной работы имеют широкие перспективы как по уточнению предложенных оценок,

так и по созданию новых исследований с результатами данной работы в качестве основы. Например, получаемые профиль компетенций злоумышленика и ресурсная база могут стать значимым подспорьем для прогнозирования и предотвращение социоинженерных атак.

Список литературы

[1] 2017 Verizon DBIR Social Engineering Breakdown. — 2017. — URL: https://www.social-engineer.com/2017-verizon-dbir-social-engineering-breakdown/ (online; accessed: 11.04.2018).

[2] 2018 Verizon Data Breach Investigations Report. — 2018. — URL: https://www.researchgate.net/profile/Suzanne_Widup/publication/324455350_2018_Verizon_Data_Breach_Investigations_Report/links/5ace9f0b0f7e9b18965a5fe5/2018-Verizon-Data-Breach-Investigations-Report.pdf?origin=publication_detail (online; accessed: 11.04.2018).

[3] Bell D. C. Trevino R. A. Modeling HIV Risk [Epidemiology] // J. Acquir Immune Defic Syndr. — 22. — 1999. — 2. — P. 280–287.

[4] Facebook готов заплатить 100 тысяч долларов за безопасность. — 2018. — URL: https://arinteg.ru/about/news/detail.php?ID=134140 (дата обращения: 11.04.2018).

[5] GetContact: найди контакт или слей контакты? — 2018. — URL: https://www.kaspersky.ru/blog/getcontact-collects-personal-data/19795/ (дата обращения: 11.04.2018).

[6] Igor Kotenko Andrey Chechulin Alexander Branitskiy. Generation of Source Data for Experiments with Network Attack Detection Software // Journal of Physics: Conference Series. — 2017. — P. 236–245.

[7] Muncaster Phil. UK Fraud Attacks Hit 20 Million in Q2. — URL: https://www.infosecurity-magazine.com/news/uk-fraud-attacks-hit-20-million-in/ (online; accessed: 11.04.2018).

[8] Wikipedia: Интегрированная среда разработки. — URL: https://ru.wikipedia.org/wiki/Интегрированная_среда_разработки (дата обращения: 11.04.2018).

[9] Абрамов М.В. Методы и алгоритмы анализа защищённости пользователей информационных систем от социоинженерных атак: оценка параметров моделей: 05.13.19 : Дисс... кандидата наук / М.В. Абрамов. — СПб., 2018. — С. 232.

[10] Абрамов М.В. Азаров А.А. Фильченков А.А. Распространение социоинженерной атаки злоумышленника на пользователей информационной системы, представленных в виде графа социальных связей // Сборник докладов Международной конференции по мягким вычислениям и измерениям (SCM-2015). — 1-2. — Санкт-Петербург, 2015. — С. 329–332.

[11] Абрамов М.В. Азаров А.А. Тулупьева Т.В. Тулупьев А.Л. Модель профиля компетенций злоумышленника в задаче анализа защищённости персонала информационных систем от социоинженерных атак // Информационно-управляющие системы. — 2016. — № 4. — С. 77–84.

[12] Азаров А.А. Тулупьев А.Л. Соловцов Н.Б. Тулупьева Т.В. Ускорение расчетов оценки защищенности пользователей информационной системы за счет элиминации маловероятных траекторий социоинженерных атак // Труды СПИИРАН. — 52. — 2013. — 2. — С. 171–181.

[13] Азаров А.А. Абрамов М.В. Тулупьева Т.В. Фильченков А.А. Применение вероятностно-реляционных моделей комплекса «критичные документы – информационная система – пользователь – злоумышленник» для анализа защищенности пользователей информационных систем от социоинженерных атак // Нечеткие системы и мягкие вычисления. — 2015. — Т. 10, № 2. — С. 209–221.

[14] Азаров А.А. Тулупьева Т.В. Суворова А.В. Тулупьев А.Л. Абрамов М.В. Юсупов Р.М. Социоинженерные атаки: проблемы анализа. — СПб. : Наука, 2016. — С. 352.

[15] Бесплатный кофе, шпионское такси и дырка в аэропорту. — 2018. — URL: https://www.kaspersky.ru/blog/small-hacks-sas2018/19912/ (дата обращения: 11.04.2018).

[16] Бушмелёв Ф.В. Абрамов М.В. Подход к построению профиля компетенций злоумышленника в задаче анализа защищённости информационной системы от социоинженерных атак // Информационная безопасность регионов России (ИБРР-2017). X Санкт-Петербургская межрегиональная конференция. (Санкт-Петербург, 1–3 ноября 2017 г.): Материалы конференции. — СПб. : СПОИСУ, 2017. — С. 414–415.

[17] Бушмелёв Ф.В. Абрамов М.В. Обзор программного инструментария для визуализации сетей в микромире корпоративных офисов // Труды VII всероссийской научно-практической конференции «Нечёткие системы, мягкие вычисления и интеллектуальные технологии» НСМВИТ–2017 (г. Санкт-Петербург, 3–7 июля, 2017 г.). — 2. — СПб. : Политехника-сервис, 2017. — С. 34–42.

[18] Искусственный нейрон. — URL: https://ru.wikipedia.org/wiki/Искусственный_нейрон (дата обращения: 11.04.2018).

[19] Кибербезопасность по-прежнему актуальна для российских компаний. — 2018. — URL: https://arinteg.ru/about/news/detail.php?ID=134153 (дата обращения: 11.04.2018).

[20] Котенко И. В. Саенко И. Б. Коцыняк М. А. Лаута О. С. Оценка киберустойчивости компьютерных сетей на основе моделирования кибератак методом преобразования стохастических сетей // Труды СПИИРАН. — 55. — 2017. — С. 160–184.

[21] Пока вирус не разлучит нас: «Лаборатория Касперского» узнала, чем может грозить совместное использование влюбленными устройств и аккаунтов. — 2018. — URL: https://www.kaspersky.ru/about/press-releases/2018_love-getting-in-the-way-of-users-internet-security (дата обращения: 11.04.2018).

[22] Специальные требования и рекомендации по технической защите конфиденциальной информации (СТР-К). — 2015. — URL: http://wiki.informationsecurity.club/doku.php/документы:нмд:стр-к (дата обращения: 11.04.2018).

[23] Утечка персональных данных: почти 2 миллиарда файлов в 2017 году. — 2018. — URL: https://arinteg.ru/about/news/detail.php?ID=134172 (дата обращения: 11.04.2018).

[24] Утечки данных в учреждениях здравоохранения: новые подробности. — 2018. — URL: https://arinteg.ru/about/news/detail.php?ID=134181 (дата обращения: 11.04.2018).

[25] ЦИК уехал — жулики остались. — 2018. — URL: https://www.kaspersky.ru/blog/russian-after-election-scam-2018/20015/ (дата обращения: 11.04.2018).

[26] Чужих уязвимостей не бывает: бреши в киберзащите партнёров обходятся бизнесу дороже всего. — 2018. — URL: https://www.kaspersky.ru/about/press-releases/2018_breaches-in-cybersecurity-of-business-partners (дата обращения: 11.04.2018).

Приложение А: список терминов

Безопасность информации — состояние защищенности информации, характеризуемое способностью персонала, технических средств и информационных технологий обеспечивать конфиденциальность (т. е. сохранение в тайне от субъектов,не имеющих полномочий на ознакомление с ней), целостность и доступность информации при ее обработке техническими средствами [22].

Граф социальных связей — граф, узлы которого представлены социальными объектами, такими как пользовательские профили с различными атрибутами (например: имя, день рождения, родной город и т.д.), сообщества, медиаконтент и т.д., а рёбра — социальными связями между ними [14].

Документ — материальный объект с зафиксированной на нем информацией в виде текста, звукозаписи или изображения, предназначенный для передачи во времени и пространстве в целях хранения и общественного использования [14].

Критичный документ — материальный объект с зафиксированной на нем информацией в виде текста, звукозаписи или изображения, имеющей ценность для компании [14].

Пользователь (потребитель) информации — субъект, обращающийся к информационной системе или посреднику за получением необходимой ему информации и пользующийся ею [14].

Интегрированная среда разработки (Integrated Development Environment, IDE) — комплекс программных средств, используемый для разработки программного обеспечения (ПО) [8].

Информационная система — организованно упорядоченная совокупность документов (массив документов) и информационных технологий, в том числе с использованием средств вычислительной техники и связи, реализующих информационные процессы. [14].

Программно-техническая атака — программно-техническое воздействие, направленное на активизацию уязвимости [11].

Профиль компетенций злоумышленника — набор компетенций

злоумышленника, характеризующих его степень владения тем или иным социоинженерным атакующих воздействием [11]. В работе формализован как совокупность пар название компетенции — степень владения атакующим воздействием.

Профиль уязвимостей пользователя — набор уязвимостей пользователя, характеризующих его склонность к тем или иным действиям в ответ на социоинженерные атакующих воздействия злоумышленника [14]. В работе формализован как совокупность пар название уязвимости — степень выраженности уязвимости.

Социоинженерная (социотехническая) атака — набор прикладных психологических и аналитических приемов, которые злоумышленники применяют для скрытой мотивации пользователей публичной или корпоративной сети к нарушениям устоявшихся правил и политик в области информационной безопасности [14].

Уязвимость пользователя — некоторая характеристика пользователя, которая делает возможным успех социоинженерного атакующего действия злоумышленника [14].